

WHAT IS CLAIMED IS:

1. A computer-implemented method for processing an electronic document, the method comprising:

obtaining a first electronic document including a plurality of logical elements, each logical element having a logical type and having associated content with a visual appearance;

generating a second electronic document by associating a marker attribute value, the marker attribute value being a value of a marker attribute, with a plurality of logical elements in the first electronic document and converting the first electronic document with the associated marker attribute values through a document conversion process that preserves the association of the marker attribute values and the logical elements; and

using the marker attribute values to identify logical elements in the second electronic document.

2. The method of claim 1, further comprising the steps of:

generating a third electronic document, before associating a marker attribute value with the plurality of logical elements in the first electronic document, by transferring the first electronic document through the document conversion process; and

using the marker attribute values in the second electronic document to identify logical elements in the third electronic document.

3. The method of claim 2, wherein each of the logical elements in the second and third electronic documents has a corresponding logical element in the first electronic document.

4. The method of claim 1, wherein each of the logical elements in the second electronic document has a corresponding logical element in the first electronic document.

5. The method of claim 1, wherein the document conversion process is a print process.

6. The method of claim 1, wherein the step of generating a second electronic document comprises associating a different marker attribute value with each of the plurality of logical elements in the first electronic document.

7. The method of claim 1, wherein the step of generating a second electronic document comprises associating a different marker attribute value with each logical element located within one same page of the first electronic document.

8. The method of claim 1, wherein the marker attribute value is a color value.

9. The method of claim 1, wherein the first electronic document is an electronic document generated in a word processing application.

10. The method of claim 1, wherein the second electronic document is a PDF document.

11. The method of claim 1, wherein the step of using the marker attribute values to identify logical elements in the second electronic document comprises:

identifying the logical elements in the second electronic document by converting the marker attribute values to logical types.

12. The method of claim 1, further comprising using the marker attribute values in the second electronic document to create a hierarchal structure for the plurality of logical elements in the second electronic document.

13. The method of claim 12, further comprising obtaining structural information from the first electronic document to create a hierarchal structure for the plurality of logical elements in the second electronic document.

14. A computer-implemented method for converting a source document including a plurality of logical elements into a PDF document, the method comprising:

producing a first PDF document from the source document;

color-coding the logical elements of the source document;

producing a color-coded, second PDF document from the color-coded source document; and

creating logical elements in the first PDF document based on the color-coded PDF document, each logical element of the color-coded PDF document corresponding to a logical element in the source document.

5 15. A computer-implemented method for creating a final-format document having logical elements from source document generated by a computer application, the method comprising:

obtaining an original final-format document from the computer application, the original final-format document being generated from the source document;

10 marking logical elements of the source document;

obtaining a marked final-format document from the computer application, the marked final-format document being generated from the marked source document;

obtaining logical structure information from the source application; and

15 creating logical elements in the original final-formatted document using the obtained logical structure information and the marked final-format document.

16. A computer program product, stored on a machine-readable medium, comprising instructions operable to cause a programmable processor to:

20 obtain a first electronic document including a plurality of logical elements, each logical element having a logical type and having associated content with a visual appearance;

25 generate a second electronic document by associating a marker attribute value, the marker attribute value being a value of a marker attribute, with a plurality of logical elements in the first electronic document and converting the first electronic document with the associated marker attribute values through a document conversion process that preserves the association of the marker attribute values and the logical elements; and

use the marker attribute values to identify logical elements in the second electronic document.

30 17. A computer program product, stored on a machine-readable medium, comprising instructions operable to cause a programmable processor to:

produce a first PDF document from the source document;

color-code the logical elements of the source document;
produce a color-coded, second PDF document from the color-coded source
document; and

5 create logical elements in the first PDF document based on the color-coded PDF
document, each logical element of the color-coded PDF document corresponding to a logical
element in the source document.

18. A computer program product, stored on a machine-readable medium, comprising
instructions operable to cause a programmable processor to:

10 obtain an original final-format document from the computer application, the original
final-format document being generated from the source document;

mark logical elements of the source document;

obtain a marked final-format document from the computer application, the marked
final-format document being generated from the marked source document;

15 obtain logical structure information from the source application; and

create logical elements in the original final-formatted document using the obtained
logical structure information and the marked final-format document.